Figure 1
Schematic layout of the arrangement of the genetic locus encoding the signal peptide precursor, the histidine kinase and the response regulator. Note that this arrangement is different from other loci in related streptococci for the following reasons: a) The comC gene is transcribed from its own promoter alone, unlike the genes thus far described in other streptococci that are arranged in an operon-like cluster with the comC/DE genes being transcribed from a single promoter.
b) The comC gene is separated from the *comD* gene by 148 nucleotides.

Streptococcus mutans
***ComCDE* Operon**

Figure 2
Sequences of the open reading frames encoding the signal peptide precursor (ComC), the histidine kinase (ComD), and the response regulator (ComE).

> S. mutans comC gene
Encodes a precursor to a signal peptide
[ATGAAAAAAACACTATCATTAAAAAATGACTTTAAAGAAATTAAGACTGATGAATTAG
AGATTATCATTGGCGGA(AGCGGAAGCCTATCAACATTTTTCCGGCTGTTTAACAGAAG
TTTTACACAAGCTTTGGGAAAA)]TAA

> S. mutans CSP encoding sequence
Competence Signal Peptide
AGCGGAAGCCTATCAACATTTTTCCGGCTGTTTAACAGAAGTTTTACACAAGCTTTGGG
AAAA [SEQ ID NO:1]

> S. mutans comD gene
Encodes a protein that functions as a histidine kinase receptor
[ATGAATGAAGCCTTAATGATACTTTCAAATGGTTTATTAACTTATCTAACCGTTCTAT
TTCTCTTGTTTCTATTTTCTAAGGTAAGTAATGTCACTTTATCGAAAAAGGAATTAACT
CTTTTTTCGATAAGCAATTTTCTGATAATGATTGCTGTTACGATGGTGAACGTAAACCT
GTTTTATCCTGCAGAGCCTCTTTATTTTATAGCTTTATCAATTTATCTTAATAGACAGA
ATAGTCTTTCTCTAAATATATTTTATGGTCTGCTGCCTGTTGCCAGTTCTGACTTGTTT
AGGCGGGCAATCATATTCTTTATCTTGGATGGAACTCAAGGAATTGTAATGGGCAGTAG
CATTATAACCACCTATATGATCGAGTTTGCAGGAATAGCGCTAAGTTACCTCTTTCTCA
GTGTGTTCAATGTTGATATTGGTCGACTTAAAGATAGTTTGACCAAGATGAAGGTCAAA
AAACGCTTGATTCCAATGAATATTACTATGCTTCTATACTACCTTTTAATACAGGTATT
GTATGTTATAGAGAGTTATAATGTGATACCGACTTTAAAATTTCGTAAATTTGTCGTTA
TTGTCTATCTTATTTTATTTTTGATTCTGATCTCATTTTTAAGCCAATATACCAAACAA
AAGGTTCAAAATGAGATAATGGCACAAAAGGAAGCTCAGATTCGAAATATCACCCAGTA
TAGTCAGCAAATAGAATCTCTTTACAAGGATATTCGAAGTTTCCGCCATGATTATCTGA
ATATTTTAACTAGCCTCAGATTAGGCATTGAAAATAAAGATTTAGCTAGTATTGAAAAG
ATTTACCATCAAATCTTAGAAAAAACAGGACATCAATTGCAGGATACCCGTTATAATAT
CGGCCATCTAGCTAATATTCAAAACGATGCTGTCAAGGGTATCTTGTCAGCAAAAATCT
TAGAAGCTCAGAATAAAAAGATTGCTGTCAATGTAGAAGTCTCAAGTAAAATACAACTG
CCTGAGATGGAGTTGCTTGATTTCATTACCATACTTTCTATCTTGTGTGATAATGCCAT
TGAGGCTGCTTTCGAATCATTAAATCCTGAAATTCAGTTAGCCTTTTTTAAGAAAAATG
GCAGTATAGTCTTTATCATTCAGAATTCCACCAAAGAAAAACAAATAGATGTGAGTAAA
ATTTTTAAAGAAAACTATTCCACTAAAGGCTCCAATCGCGGTATTGGTTTAGCAAAGGT
GAATCATATTCTTGAACATTATCCCAAAACCAGTTTACAAACAAGCAATCATCATCATT
TATTCAAGCAACTCCTAATAATAAAA]TAG

> S. mutans comE gene
Encodes a response regulator that activates transcription of a number of genes

[ATGATTTCTATTTTTGTATTGGAAGATGATTTTTTACAACAAGGACGTCTTGAAACCA
CCATTGCAGCTATCATGAAAGAAAAAAATTGGTCTTATAAAGAATTGACTATTTTTGGA
AAACCACAACAACTTATTGACGCTATCCCTGAAAAGGGCAATCACCAGATTTTCTTTTT
GGATATTGAAATCAAAAAAGAGGAAAAGAAAGGACTGGAAGTAGCCAATCAGATTAGAC
AGCATAATCCTAGTGCAGTTATTGTCTTTGTCACGACACATTCTGAGTTTATGCCCCTC
ACTTTTCAGTATCAGGTATCTGCTTTGGATTTTATTGATAAATCTTTGAATCCTGAGGA
GTTCTCCCACCGCATTGAATCAGCGCTGTATTATGCTATGGAAAACAGCCAGAAGAATG
GTCAATCAGAGGAACTTTTTATTTTCCATTCATCTGAAACTCAGTTTCAGGTCCCTTTT
GCTGAGATTCTGTATTTTGAAACATCTTCAACAGCCCATAAGCTCTGCCTTTATACTTA
TGATGAACGGATTGAATTCTACGGCAGTATGACTGACATTGTTAAAATGGATAAGAGAC
TTTTTCAGTGCCATCGCTCTTTTATTGTCAATCCTGCCAATATTACCCGTATTGATCGG
AAAAAACGCTTGGCCTATTTTCGAAATAATAAGTCTTGTCTTATTTCACGAACTAAGTT
AACAAAACTGAGAGCTGTGATTGCTGATCAAAGGAGAGCAAAA] TGA

Figure 3
The amino acid sequences of the signal peptide precursor (ComC), the histidine kinase (ComD), and the response regulator (ComE).

> S. mutans ComC protein (CSP Precursor)
MKKTLSLKNDFKEIKTDELEIIIGGSGSLSTFFRLFNRSFTQALGK

> S. mutans ComD protein (Histidine Kinase)
MNEALMILSNGLLTYLTVLFLLFLFSKVSNVTLSKKELTLFSISNFLIMIAVTMVNVNL
FYPAEPLYFIALSIYLNRQNSLSLNIFYGLLPVASSDLFRRAIIFFILDGTQGIVMGSS
IITTYMIEFAGIALSYLFLSVFNVDIGRLKDSLTKMKVKKRLIPMNITMLLYYLLIQVL
YVIESYNVIPTLKFRKFVVIVYLILFLILISFLSQYTKQKVQNEIMAQKEAQIRNITQY
SQQIESLYKDIRSFRHDYLNILTSLRLGIENKDLASIEKIYHQILEKTGHQLQDTRYNI
GHLANIQNDAVKGILSAKILEAQNKKIAVNVEVSSKIQLPEMELLDFITILSILCDNAI
EAAFESLNPEIQLAFFKKNGSIVFIIQNSTKEKQIDVSKIFKENYSTKGSNRGIGLAKV
NHILEHYPKTSLQTSNHHHLFKQLLIIK

> S. mutans ComE protein (Response Regulator)
MISIFVLEDDFLQQGRLETTIAAIMKEKNWSYKELTIFGKPQQLIDAIPEKGNHQIFFL
DIEIKKEEKKGLEVANQIRQHNPSAVIVFVTTHSEFMPLTFQYQVSALDFIDKSLNPEE
FSHRIESALYYAMENSQKNGQSEELFIFHSSETQFQVPFAEILYFETSSTAHKLCLYTY
DERIEFYGSMTDIVKMDKRLFQCHRSFIVNPANITRIDRKKRLAYFRNNKSCLISRTKL
TKLRAVIADQRRAK

Figure 4
The deduced amino acid sequence of the signal peptide precursor in various strains and its predicted cleavage site. The original peptide is expressed as a 46-amino acid peptide that is cleaved after the glycine-glycine residues to generate an active signal peptide.

```
BM71 CSP      1 MKKTPSLKNDFKEIKTDELEIIIGGSGSLSTFFRLFNRSFTQALGK 46
GB14 CSP      1 MKKTLSLKNDFKEIKTDELEIIIGGSGSLSTFFRLFNRSFTQALGK 46
H7 CSP        1 MKKTLSLKNDFKEIKTDELEIIIGGSGSLSTFFRLFNRSFTQALGK 46
JH1005 CSP    1 MKKTLSLKNDFKEIKTDELEIIIGGSGTLSTFFRLFNRSFTQA    43
LT11 CSP      1 MKKTLSLKNDFKEIKTDELEIIIGGSGSLSTFFRLFNRSFTQALGK 46
NG8 CSP       1 MKKTLSLKNDFKEIKTDELEIIIGGSGSLSTFFRLFNRSFTQALGK 46
UAB159 CSP    1 MKKTLSLKNDFKEIKTDELEIIIGGSGSLSTFFRLFNRSFTQALGK 46
                **** **********************.***************

consensus: 1 MKKTLSLKNDFKEIKTDELEIIIGG SGSLSTFFRLFNRSFTQALGK 46
predicted cleavage site:              ^
```

Figure 5
The synthetic signal peptide that is effective at inducing competence, biofilm formation and acid tolerance in *Streptococcus mutans*.

```
SGSLSTFFRLFNRSFTQALGK [SEQ ID NO:2]
```

Figure 6
The natural activity of the signal/receptor system functioning *in vitro* in model biofilms as determined by the ability of various strains of *S. mutans* to accept donor plasmid DNA conferring erythromycin resistance.

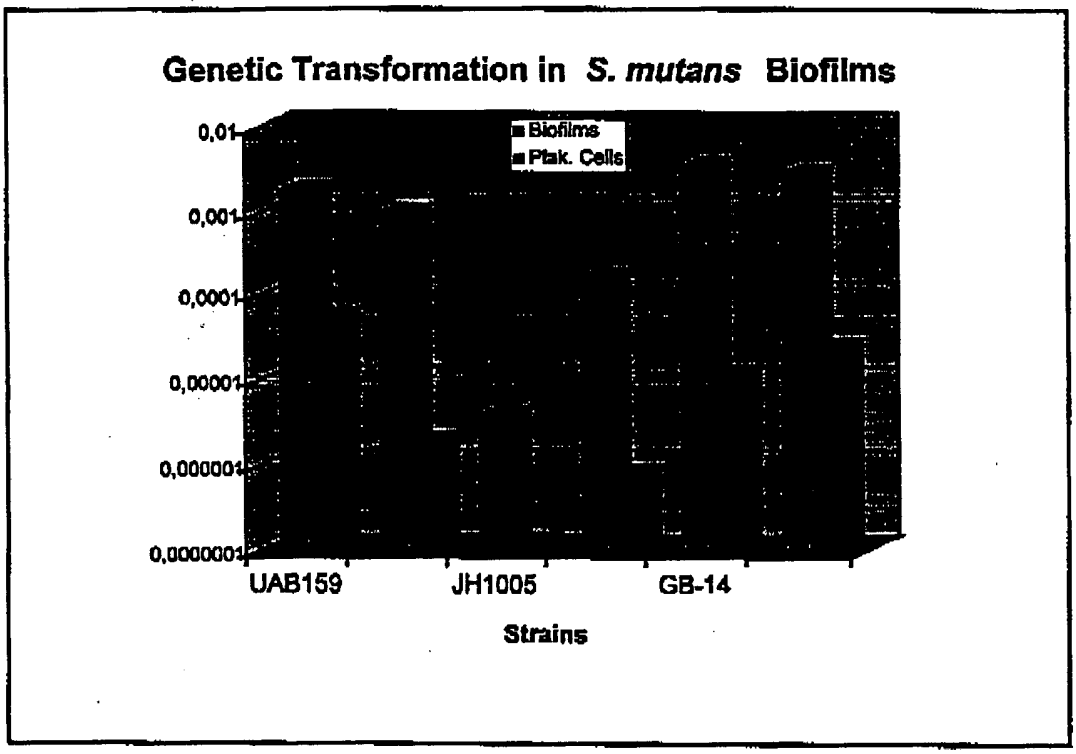


Figure 7
Induction of genetic transformation in *Streptococcus mutans* by synthetic competence stimulating peptide (SCSP)[1]

Strain	Peptide added Number of Transformants/Recipients	No peptide Number of Transformants/Recipients
UAB15	4.65×10^{-1}	1.78×10^{-6}
JH1005[2]	6.98×10^{-2}	0

[1]The final concentration of SCSP used was 500 ng/ml.
The strain contains a nonsense mutation in the *comC* gene encoding the CSP.

Figure 8
List of the primers used to amplify the genes or internal regions of the target genes by polymerase chain reaction (PCR) for subsequent sequencing or inactivation.

ComC region

ComC Primer Pair: F5-B5

[F5] 23406-23424 5'- AGTTTTTTGTCTGGCTGCG -3'
19 nt forward primer
pct G+C: 47.4 Tm: 50.5

[B5] 24056-24037 5'- TCCACTAAAGGCTCCAATCG -3'
20 nt backward primer
pct G+C: 50.0 Tm: 51.9

651 nt product for F5-B5 pair (23406-24056)
Optimal annealing temp: 50.3
pct G+C: 30.9 Tm: 71.5

ComD region

ComD Primer Pair: F1-B1

[F1] 392-415 5'- CGCTAAGTTACCTCTTTCTCAGTG -3'
24 nt forward primer
pct G+C: 45.8 Tm: 51.6

[B1] 683-663 5'- GCTTCCTTTTGTGCCATTATC -3'
21 nt backward primer
pct G+C: 42.9 Tm: 50.8

292 nt product for F1-B1 pair (392-683)
Optimal annealing temp: 49.5
pct G+C: 30.8 Tm: 70.2

ComE region

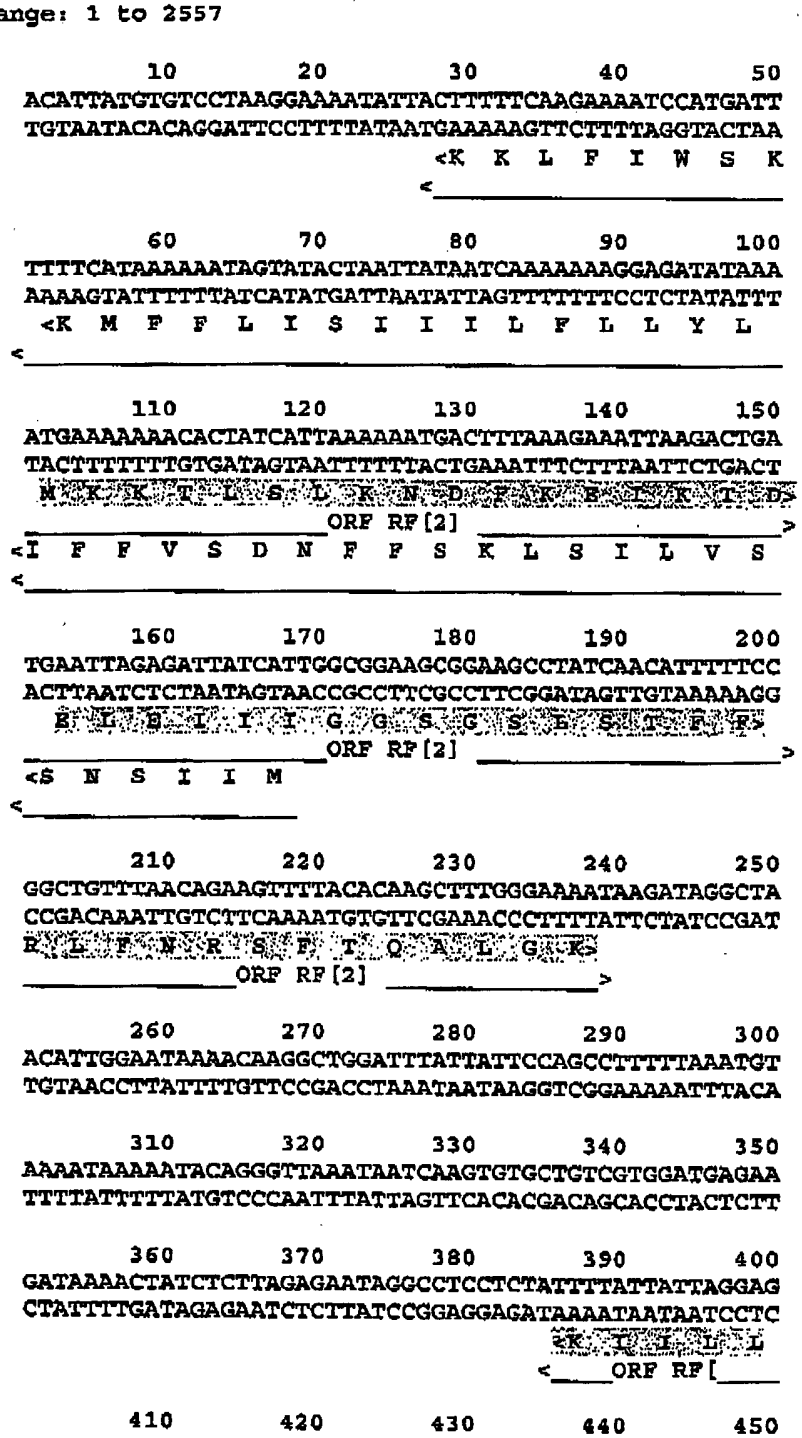
ComE Primer Pair: F1-B1

[F1] 145-165 5'- CCTGAAAAGGGCAATCACCAG -3'
21 nt forward primer
pct G+C: 52.4 Tm: 55.9

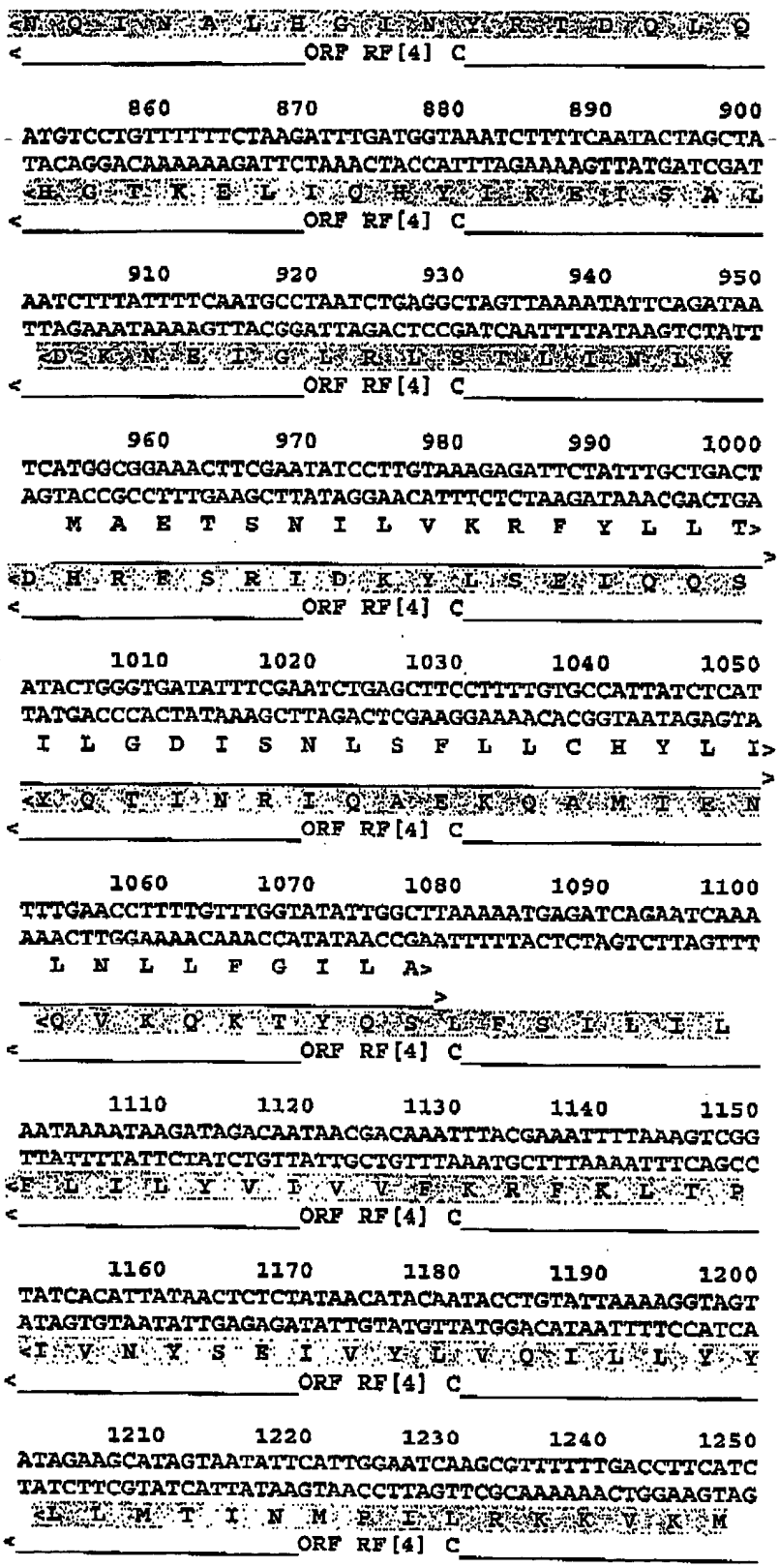
[B1] 606-585 5'- GCGATGGCACTGAAAAAGTCTC -3'
22 nt backward primer
pct G+C: 50.0 Tm: 55.4

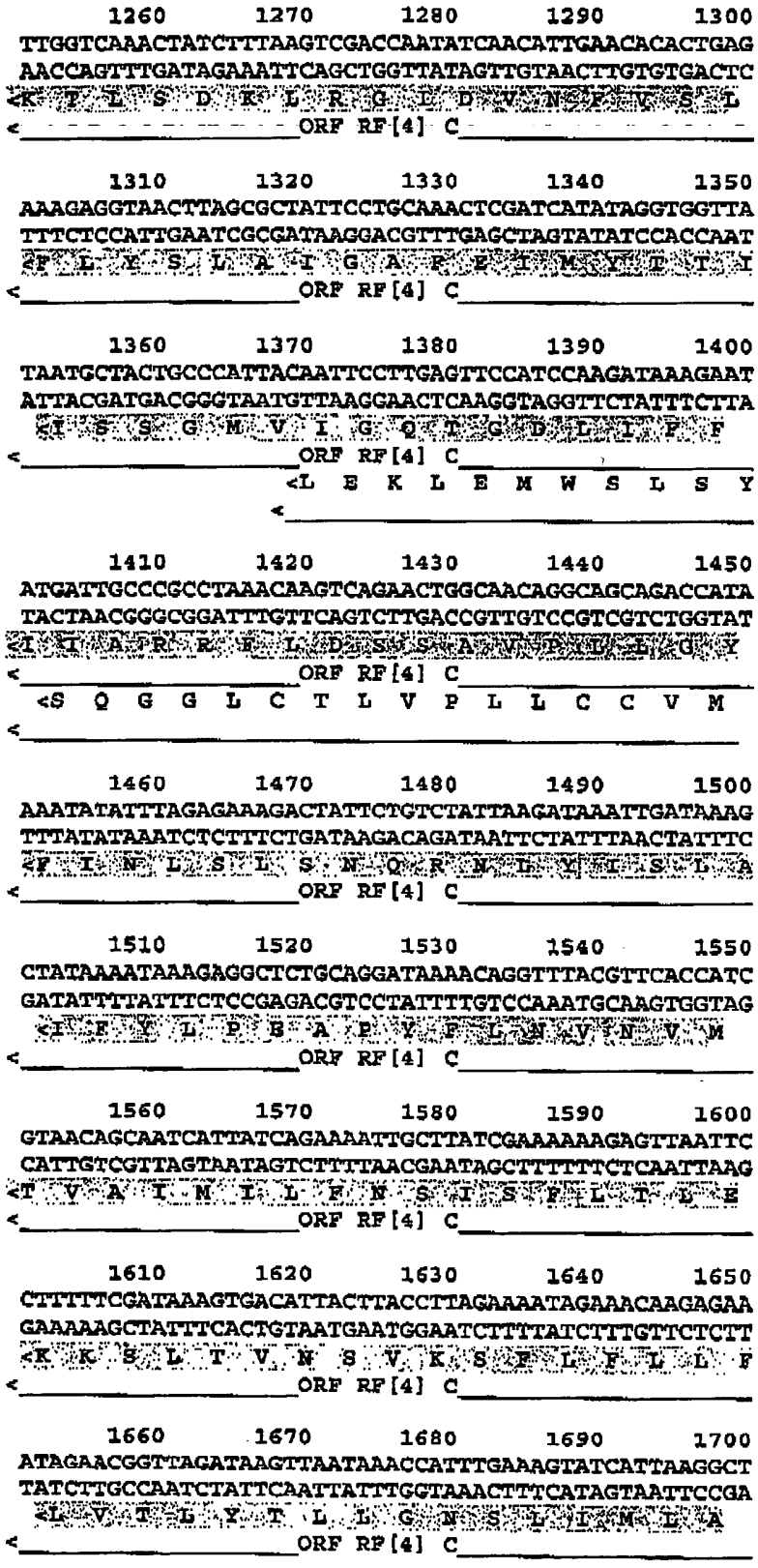
462 nt product for F1-B1 pair (145-606)
Optimal annealing temp: 53.6
pct G+C: 38.3 Tm: 74.1

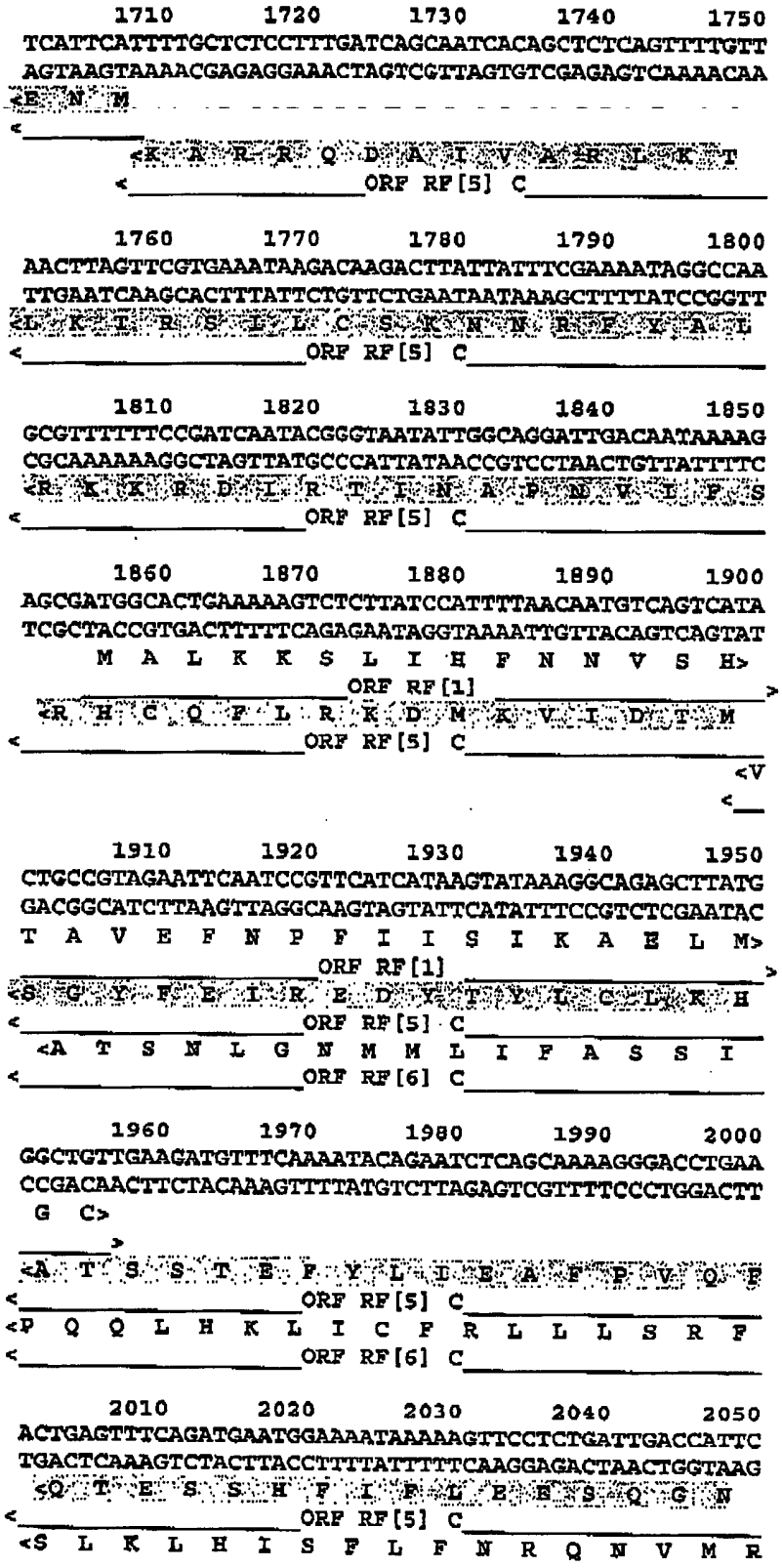
Figure 9
ComCDE local region. The ComC (first highlighted region; nucleotides 101 to 241), ComD (second highlighted region; nucleotides 383 to 1708) and ComE (third highlighted region; nucleotides 1705 to 2457) proteins are highlighted.



```
AACGAACTTATTTACTACTACTAACGAACAAACATTTGACCAAAACCCTA
<Q  K  F  L  H  H  H  N  S  T  Q  L  S  T  K  P  Y
<________________________ORF RF[4] C__________________________

        460       470       480       490       500
AATGTTCAAGAATATGATTCACCTTTGCTAAACCAATACCGCGATTGGAG
TTACAAGTTCTTATACTAAGTGGAAACGATTTGGTTATGGCGCTAACCTC
 <H  E  L  I  H  N  V  K  A  L  G  I  G  R  N  S
<________________________ORF RF[4] C__________________________

        510       520       530       540       550
CCTTTAGTGGAATAGTTTTCTTTAAAAATTTTACTCACATCTATTTGTTT
GGAAATCACCTTATCAAAAGAAATTTTTAAAATGAGTGTAGATAAACAAA
<G  K  T  S  Y  N  E  K  F  I  K  S  V  D  I  Q  K
<________________________ORF RF[4] C__________________________

        560       570       580       590       600
TTCTTTGGTGGAATTCTGAATGATAAAGACTATACTGCCATTTTTCTTAA
AAGAAACCACCTTAAGACTTACTATTTCTGATATGACGGTAAAAAGAATT
<E  K  T  S  N  Q  I  I  F  V  I  S  G  N  K  K  F
<________________________ORF RF[4] C__________________________

        610       620       630       640       650
AAAAGGCTAACTGAATTTCAGGATTTAATGATTCGAAAGCAGCCTCAATG
TTTTCCGATTGACTTAAAGTCCTAAATTACTAAGCTTTCGTCGGAGTTAC
                                                M>
                                              ____>
 <F  A  L  Q  I  E  P  N  L  S  E  F  A  A  E  I
<________________________ORF RF[4] C__________________________

        660       670       680       690       700
GCATTATCACACAAGATAGAAAGTATGGTAATGAAATCAAGCAACTCCAT
CGTAATAGTGTGTTCTATCTTTCATACCATTACTTTAGTTCGTTGAGGTA
 A  L  S  H  K  I  E  S  M  V  M  K  S  S  N  S  I>
_________________________ORF RF[3]____________________________>
<A  N  D  C  L  I  S  L  T  T  I  F  D  L  L  E  M
<________________________ORF RF[4] C__________________________

        710       720       730       740       750
CTCAGGCAGTTGTATTTTACTTGAGACTTCTACATTGACAGCAATCTTTT
GAGTCCGTCAACATAAAATGAACTCTGAAGATGTAACTGTCGTTAGAAAA
  S  G  S  C  I  L  L  E  T  S  T  L  T  A  I  F>
_________________________ORF RF[3]____________________________>
<E  P  L  Q  I  K  S  S  V  E  V  N  V  A  I  K  K
<________________________ORF RF[4] C__________________________

        760       770       780       790       800
TATTCTGAGCTTCTAAGATTTTTGCTGACAAGATACCCTTGACAGCATCG
ATAAGACTCGAAGATTCTAAAAACGACTGTTCTATGGGAACTGTCGTAGC
L  F>
______>
 <N  Q  A  E  L  I  K  A  S  L  I  G  K  V  A  D
<________________________ORF RF[4] C__________________________

        810       820       830       840       850
TTTTGAATATTAGCTAGATGGCCGATATTATAACGGGTATCCTGCAATTG
AAAACTTATAATCGATCTACCGGCTATAATATTGCCCATAGGACGTTAAC
```


ORF RF[4]
860 870 880 890 900
910 920 930 940 950
960 970 980 990 1000
1010 1020 1030 1040 1050
1060 1070 1080 1090 1100
1110 1120 1130 1140 1150
1160 1170 1180 1190 1200
1210 1220 1230 1240 1250





```
<____________________ORF RF[6] C____________________

        2060      2070      2080      2090      2100
TTCTGGCTGTTTTCCATAGCATAATACAGCGCTGATTCAATGCGGTGGGA
AAGACCGACAAAAGGTATCGTATTATGTCGCGACTAAGTTACGCCACCCT
<K  Q  S  N  E  M  A  Y  Y  L  A  S  E  I  R  H  S
<____________________ORF RF[5] C____________________
  <R  A  T  K  W  L  M
<_____ORF RF[6] C_______

        2110      2120      2130      2140      2150
GAACTCCTCAGGATTCAAAGATTTATCAATAAAATCCAAAGCAGATACCT
CTTGAGGAGTCCTAAGTTTCTAAATAGTTATTTTAGGTTTCGTCTATGGA
<F  E  E  P  N  L  S  K  D  I  F  D  K  A  S  V  Q
<____________________ORF RF[5] C____________________

        2160      2170      2180      2190      2200
GATACTGAAAAGTGAGGGGCATAAACTCAGAATGTGTCGTGACAAAGACA
CTATGACTTTTCACTCCCCGTATTTGAGTCTTACACAGCACTGTTTCTGT
                                  M  C  R  D  K  D>
                                  ________________>
 <Y  Q  F  T  L  P  M  F  E  S  H  T  T  V  F  V
<____________________ORF RF[5] C____________________

        2210      2220      2230      2240      2250
ATAACTGCACTAGGATTATGCTGTCTAATCTGATTGGCTACTTCCAGTCC
TATTGACGTGATCCTAATACGACAGATTAGACTAACCGATGAAGGTCAGG
N  N  C  T  R  I  M  L  S  N  L  I  G  Y  F  Q  S>
__________________________________________________>
<I  V  A  S  D  N  H  Q  R  I  Q  N  A  V  E  D  G
<____________________ORF RF[5] C____________________

        2260      2270      2280      2290      2300
TTTCTTTTCCTCTTTTTTTGATTTCAATATCCAAAAAGAAAATCTGGTGAT
AAAGAAAAGGAGAAAAAAACTAAAGTTATAGGTTTTTCTTTTAGACCACTA
 F  L  F  L  P  F  D  F  N  I  Q  K  E  N  L  V  I>
__________________________________________________>
<K  K  E  E  K  K  I  E  I  D  L  F  F  I  Q  H  N
<____________________ORF RF[5] C____________________

        2310      2320      2330      2340      2350
TGCCCTTTTCAGGGATAGCGTCAATAAGTTGTTGTGGTTTTCCAAAAATA
ACGGGAAAAGTCCCTATCGCAGTTATTCAACAACACCAAAAGGTTTTTAT
 A  L  F  R  D  S  V  N  K  L  L  W  F  S  K  N>
__________________________________________________>
 <G  K  E  P  I  A  D  I  L  Q  Q  P  K  G  F  I
<____________________ORF RF[5] C____________________

        2360      2370      2380      2390      2400
GTCAATTCTTTATAAGACCAATTTTTTTCTTTCATGATAGCTGCAATGGT
CAGTTAAGAAATATTCTGGTTAAAAAAAGAAAGTACTATCGACGTTACCA
S  Q  F  F  I  R  P  I  F  F  F  H  D  S  C  N  G>
__________________________________________________>
                                   M  I  A  A  M  V>
                                   _______________>
<T  L  E  K  Y  S  W  N  K  E  K  M  I  A  A  I  T
<____________________ORF RF[5] C____________________
```

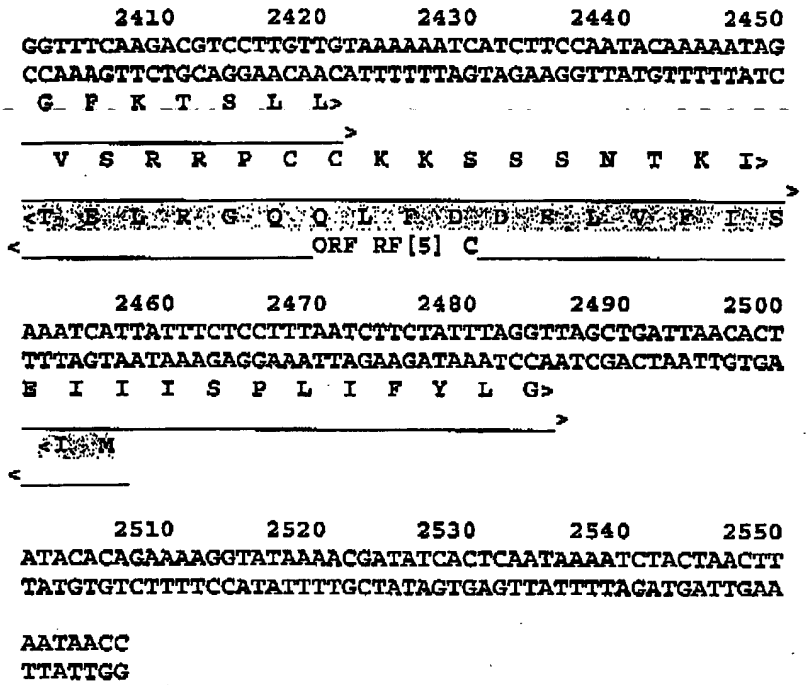


Figure 10
The comX nucleotide sequence, amino acid sequence, and its local region with 100bp included both upstream and downstream (promoter is upstream).

> S. mutans comX gene

ATGGAAGAAGATTTTGAAATTGTTTTTAATAAGGTTAAGCCAATTGTATGGAAATTAAG
CCGTTATTACTTTATTAAAATGTGGACTCGTGAAGATTGGCAACAAGAGGGAATGTTGA
TTTTGCACCAATTATTAAGGGAACATCCAGAATTAGAAGAGGATGATACAAAATTGTAT
ATCTATTTTAAGACACGTTTTTCTAATTACATTAAAGATGTTTTGCGTCAGCAAGAAAG
TCAGAAACGTCGTTTTAATAGAATGTCTTATGAAGAAGTCGGTGAGATTGAACACTGTT
TGTCAAGTGGCGGTATGCAATTGGATGAATATATTTTATTTCGTGATAGTTTGCTTGCA
TATAAACAAGGTCTGAGTACTGAAAAGCAAGAGCTGTTTGAGCGCTTGGTAGCAGGAGA
GCACTTTTTGGGAAGGCAAAGTATGCTGAAAGATTTACGTAAAAAATTAAGTGATTTTA
AGGAAAAA

> S. mutans ComX protein

MEEDFEIVFNKVKPIVWKLSRYYFIKMWTREDWQQEGMLILHQLLREHPELEEDDTKLY
IYFKTRFSNYIKDVLRQQESQKRRFNRMSYEEVGEIEHCLSSGGMQLDEYILFRDSLLA
YKQGLSTEKQELFERLVAGEHFLGRQSMLKDLRKKLSDFKEK

> S. mutans comX gene local region

GTAAATAAAACAGCCAGTTAAGATGGGACATTTATGTCCTGTTCTTAAAGTCTTTTTCG
TTTTATAATAATTTTATTATAAAAGGAGGTCATCGTAATAGATGGAAGAAGATTTTGAA
ATTGTTTTTAATAAGGTTAAGCCAATTGTATGGAAATTAAGCCGTTATTACTTTATTAA
AATGTGGACTCGTGAAGATTGGCAACAAGAGGGAATGTTGATTTTGCACCAATTATTAA
GGGAACATCCAGAATTAGAAGAGGATGATACAAAATTGTATATCTATTTTAAGACACGT
TTTTCTAATTACATTAAAGATGTTTTGCGTCAGCAAGAAAGTCAGAAACGTCGTTTTAA
TAGAATGTCTTATGAAGAAGTCGGTGAGATTGAACACTGTTTGTCAAGTGGCGGTATGC
AATTGGATGAATATATTTTATTTCGTGATAGTTTGCTTGCATATAAACAAGGTCTGAGT
ACTGAAAAGCAAGAGCTGTTTGAGCGCTTGGTAGCAGGAGAGCACTTTTTGGGAAGGCA
AAGTATGCTGAAAGATTTACGTAAAAAATTAAGTGATTTTAAGGAAAAATAGTTAAAAA
GGGAAAGAATGGAACATGTGATTGTACCATTCTTTTTGGTTGAAAATTAAGAAAAGTTA
TTATAAATTATTGGTTTAACATGCCATATTA

Figure 11.
The comA and comB nucleotide and amino acid sequences. ComA and ComB are the components of the CSP exporter.

➢ S. mutans comA gene

ATGAAACAAGTTATTTATGTTGTTTTAATCGTCATAGCCGTTAACATTCTCTTAGAGAT
TATCAAAAGAGTAACAAAAAGGGGAGGGACAGTTTCGTCATCTAATCCTTTACCAGATG
GGCAGTCTAAGTTGTTTTGGCGCAGACATTATAAGCTAGTACCTCAGATTGATACCAGA
GACTGTGGGCCGGCAGTGCTGGCATCTGTTGCAAAGCATTACGGATCTAATTACTCTAT
CGCTTATCTGCGGGAACTCTCAAAGACTAACAAGCAGGGAACAACAGCTCTTGGCATTG
TTGAAGCTGCTAAAAAGTTAGGCTTTGAAACACGCTCTATCAAGGCGGATATGACGCTT
TTTGATTATAATGATTTGACCTATCCTTTTATCGTCCATGTGATTAAAGGAAAACGTCT
GCAGCATTATTATGTCGTCTATGGCAGCCAGAATAATCAGCTGATTATTGGAGATCCTG
ATCCTTCAGTTAAGGTGACTAGGATGAGTAAGGAACGCTTTCAATCAGAGTGGACAGGC
CTTGCAATTTTCCTAGCTCCTCAGCCTAACTATAAGCCTCATAAAGGTGAAAAAAATGG
TTTGTCTAATTTCTTCCCGTTGATCTTTAAGCAGAAAGCTTTGATGACTTATATTATCA
TAGCTAGCTTGATTGTGACGCTCATTGATATTGTCGGATCATACTATCTCCAAGGAATA
TTGGACGAGTACATTCCTGATCAGCTGATTTCAACTTTAGGAATGATTACGATTGGTCT
GATAATAACCTATATTATCCAGCAGGTCATGGCTTTTGCAAAAGAATACCTCTTGGCCG
TACTCAGTTTGCGTTTAGTCATTGATGTTATCCTGTCTTATATCAAACATATTTTTACG
CTTCCTATGTCTTTCTTTGCGACAAGGCGAACAGGAGAAATCACGTCTCGTTTTACAGA
TGCCAATCAGATTATTGATGCTGTAGCGTCAACCATCTTTTCAATCTTTTTAGATATGA
CTATGGTAATTTTGGTTGGTGGGGTTTTGTTGGCGCAAAACAATAACCTTTTCTTTCTA
ACCTTGCTCTCCATTCCGATTTATGCCATCATTATTTTTGCTTTCTTGAAACCCTTTGA
GAAAATGAATCACGAAGTGATGGAAAGCAATGCTGTGGTAAGTTCTTCTATCATTGAAG
ATATCAATGGGATGGAAACCATTAAATCACTCACAAGTGAGTCCGCTCGTTATCAAAAC
ATTGATAGTGAATTTGTTGATTATTTGGAGAAAAACTTTAAGCTACACAAGTATAGTGC
CATTCAAACCGCATTAAAAAGCGGTGCTAAGCTTATCCTCAATGTTGTCATTCTCTGGT
ATGGCTCTCGTCTAGTTATGGATAATAAAATCTCAGTTGGTCAGCTTATCACCTTTAAT
GCTTTGCTGTCTTATTTCTCAAATCCAATTGAAAATATTATCAATCTGCAATCCAAACT
GCAGTCAGCTCGCGTTGCCAATACACGTCTTAATGAGGTCTATCTTGTCGAATCTGAAT
TTGAAAAAGACGGCGATTTATCAGAAAATAGCTTTTTAGATGGTGATATTTCGTTTGAA
AATCTTTCTTATAAATATGGATTTGGGCGAGATACCTTATCAGATATTAATTTATCAAT
CAAAAAAGGCTCCAAGGTCAGTCTAGTTGGAGCCAGTGGTTCTGGTAAAACAACTTTGG
CTAAACTGATTGTCAATTTCTACGAGCCTAACAAGGGGATTGTTCGAATCAATGGCAAT
GATTTAAAAGTTATTGATAAGACAGCTTTGCGGCGGCATATTAGCTATTTGCCGCAACA
GGCCTATGTTTTTAGTGGCTCTATTATGGATAATCTCGTTTTAGGAGCTAAAGAAGGAA
CGAGTCAGGAAGACATTATTCGTGCTTGTGAAATTGCTGAAATCCGCTCGGACATTGAA
CAAATGCCTCAGGGCTATCAGACAGAGTTATCAGATGGTGCCGGTATTTCTGGCGGTCA
AAAACAGCGGATTGCTTTAGCTAGGGCCTTATTAACACAGGCACCGGTTTTGATTCTGG
ATGAAGCCACCAGCAGTCTTGATATTTTGACAGAAAAGAAAATTATCAGCAATCTCTTA
CAGATGACGGAGAAAACAATAATTTTTGTTGCCCACCGCTTAAGCATTTCACAGCGTAC
TGACGAAGTCATTGTCATGGATCAGGGAAAAATTGTTGAACAAGGCACTCATAAGGAAC
TTTTAGCTAAGCAAGGTTTCTATTATAACCTGTTTAAT

➢ S. mutans ComA protein

MKQVIYVVLIVIAVNILLEIIKRVTKRGGTVSSSNPLPDGQSKLFWRRHYKLVPQIDTR
DCGPAVLASVAKHYGSNYSIAYLRELSKTNKQGTTALGIVEAAKKLGFETRSIKADMTL
FDYNDLTYPFIVHVIKGKRLQHYYVVYGSQNNQLIIGDPDPSVKVTRMSKERFQSEWTG
LAIFLAPQPNYKPHKGEKNGLSNFFPLIFKQKALMTYIIIASLIVTLIDIVGSYYLQGI
LDEYIPDQLISTLGMITIGLIITYIIQQVMAFAKEYLLAVLSLRLVIDVILSYIKHIFT
LPMSFFATRRTGEITSRFTDANQIIDAVASTIFSIFLDMTMVILVGGVLLAQNNNLFFL
TLLSIPIYAIIIFAFLKPFEKMNHEVMESNAVVSSSIIEDINGMETIKSLTSESARYQN
IDSEFVDYLEKNFKLHKYSAIQTALKSGAKLILNVVILWYGSRLVMDNKISVGQLITFN
ALLSYFSNPIENIINLQSKLQSARVANTRLNEVYLVESEFEKDGDLSENSFLDGDISFE
NLSYKYGFGRDTLSDINLSIKKGSKVSLVGASGSGKTTLAKLIVNFYEPNKGIVRINGN
DLKVIDKTALRRHISYLPQQAYVFSGSIMDNLVLGAKEGTSQEDIIRACEIAEIRSDIE
QMPQGYQTELSDGAGISGGQKQRIALARALLTQAPVLILDEATSSLDILTEKKIISNLL
QMTEKTIIFVAHRLSISQRTDEVIVMDQGKIVEQGTHKELLAKQGFYYNLFN

> S. mutans comB gene

ATGGATCCTAAATTTTTACAAAGTGCAGAATTTTATAGGAGACGCTATCATAATTTTGC
GACACTATTAATTGTTCCTTTGGTCTGCTTGATTATCTTCTTGGTCATATTCCTTTGTT
TTGCTAAAAAAGAAATTACAGTGATTTCTACTGGTGAAGTTGCACCAACAAAGGTTGTA
GATGTTATCCAATCTTACAGTGACAGTTCAATCATTAAAAATAATTTAGATAATAATGC
AGCTGTTGAGAAGGGAGACGTTTTAATTGAATATTCAGAAAATGCCAGTCCAAACCGTC
AGACTGAACAAAAGAATATTATAAAAGAAAGACAAAAACGAGAAGAGAAGGAAAAGAAA
AAACACCAAAAGAGCAAGAAAAAGAAGAAGTCTAAGAGCAAGAAAGCTTCCAAAGATAA
GAAAAAGAAATCGAAAGACAAGGAAAGCAGCTCTGACGATGAAAATGAGACAAAAAAGG
TTTCGATTTTTGCTTCAGAAGATGGTATTATTCATACCAATCCCAAATATGATGGTGCC
AATATTATTCCGAAGCAAACCGAGATTGCTCAAATCTATCCTGATATTCAAAAAACAAG
AAAAGTGTTAATCACCTATTATGCTTCTTCTGATGATGTTGTTTCTATGAAAAAGGGGC
AAACCGCTCGTCTTTCCTTGGAAAAAAAGGGAAATGACAAGGTTGTTATTGAAGGAAAA
ATTAACAATGTCGCTTCATCAGCAACTACTACTAAAAAAGGAAATCTCTTTAAGGTTAC
TGCCAAAGTAAAGGTTTCTAAGAAAAATAGCAAACTCATCAAGTATGGTATGACAGGCA
AGACAGTCACTGTCATTGATAAAAAGACTTATTTTGATTATTTCAAAGATAAATTACTG
CATAAAATGGATAAT

> S. mutans ComB protein

MDPKFLQSAEFYRRRYHNFATLLIVPLVCLIIFLVIFLCFAKKEITVISTGEVAPTKVV
DVIQSYSDSSIIKNNLDNNAAVEKGDVLIEYSENASPNRQTEQKNIIKERQKREEKEKK
KHQKSKKKKKSKSKKASKDKKKKSKDKESSSDDENETKKVSIFASEDGIIHTNPKYDGA
NIIPKQTEIAQIYPDIQKTRKVLITYYASSDDVVSMKKGQTARLSLEKKGNDKVVIEGK
INNVASSATTTKKGNLFKVTAKVKVSKKNSKLIKYGMTGKTVTVIDKKTYFDYFKDKLL
HKMDN

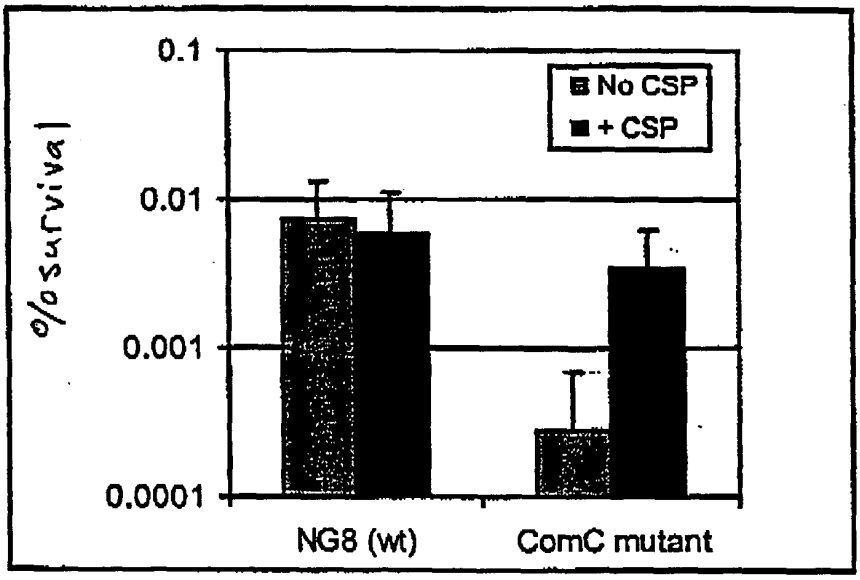


Figure 12